Linear Algebra and its Applications 435 (2011) 560–577

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Implicitly-weighted total least squares☆

Sungwoo Park [a], Dianne P. O'Leary [b,*]

[a] *Computer Science Department, University of Maryland, College Park, MD 20742, USA*
[b] *Computer Science Department and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742, USA*

ARTICLE INFO

Article history:
Available online 7 July 2010

Submitted by V. Mehrmann

Dedicated to G.W. Stewart on the occasion of his 70th birthday.

Keywords:
Least squares
Data least squares
Total least squares
Errors in variables
Linear regression
Variance estimation

ABSTRACT

In a total least squares (TLS) problem, we estimate an optimal set of model parameters $\boldsymbol{X}$, so that $(\boldsymbol{A} - \Delta\boldsymbol{A})\boldsymbol{X} = \boldsymbol{B} - \Delta\boldsymbol{B}$, where $\boldsymbol{A}$ is the model matrix, $\boldsymbol{B}$ is the observed data, and $\Delta\boldsymbol{A}$ and $\Delta\boldsymbol{B}$ are corresponding corrections. When $\boldsymbol{B}$ is a single vector, Rao (1997) and Paige and Strakoš (2002) suggested formulating standard least squares problems, for which $\Delta\boldsymbol{A} = \boldsymbol{0}$, and data least squares problems, for which $\Delta\boldsymbol{B} = \boldsymbol{0}$, as weighted and scaled TLS problems. In this work we define an implicitly-weighted TLS formulation (ITLS) that reparameterizes these formulations to make computation easier. We derive asymptotic properties of the estimates as the number of rows in the problem approaches infinity, handling the rank-deficient case as well. We discuss the role of the ratio between the variances of errors in $\boldsymbol{A}$ and $\boldsymbol{B}$ in choosing an appropriate parameter in ITLS. We also propose methods for computing the family of solutions efficiently and for choosing the appropriate solution if the ratio of variances is unknown. We provide experimental results on the usefulness of the ITLS family of solutions.

1. Introduction

In formulating a linear model $\boldsymbol{AX} \approx \boldsymbol{B}$, there can be errors in the data $\boldsymbol{B}$, errors in the model matrix $\boldsymbol{A}$, or errors in both $\boldsymbol{B}$ and $\boldsymbol{A}$. This has led to the formulation of three distinct problems: given $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{m\times d}$, where usually $m > n$, find $\boldsymbol{X}$ and small correction matrices $\Delta\boldsymbol{A}$, and $\Delta\boldsymbol{B}$ satisfying

☆ This work was supported in part by the US National Science Foundation under Grant CCF 05 14213.

* Corresponding author.
E-mail addresses: swpark81@gmail.com (S. Park), oleary@cs.umd.edu (D.P. O'Leary).

0024-3795/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2010.06.020

$$(\boldsymbol{A} - \Delta\boldsymbol{A})\boldsymbol{X} = \boldsymbol{B} - \Delta\boldsymbol{B}, \tag{1.1}$$

where

- $\Delta\boldsymbol{A} = \boldsymbol{0}$ for the *least squares* (LS) problem.
- $\Delta\boldsymbol{B} = \boldsymbol{0}$ for the *data least squares* (DLS) problem.
- Both $\Delta\boldsymbol{A}$ and $\Delta\boldsymbol{B}$ are allowed to be nonzero for the *total least squares* (TLS) problem.

In least squares formulations, the values of $\boldsymbol{X}$, $\Delta\boldsymbol{A}$, and $\Delta\boldsymbol{B}$ are found by minimizing

$$\|[\Delta\boldsymbol{A}, \Delta\boldsymbol{B}]\|_F, \tag{1.2}$$

where $\|.\|_F$ denotes the *Frobenius norm*, the square root of the sum of the squares of the entries.

Minimizing (1.2) makes sense, for example, if the errors in $\boldsymbol{A}$ and $\boldsymbol{B}$ are zero-mean, mutually uncorrelated, and drawn from the same distribution. If, on the other hand, the standard deviation of the errors in $\boldsymbol{A}$ is γ times the standard deviation of the errors in $\boldsymbol{B}$, then we should weight the terms in (1.2) as

$$\|[\Delta\boldsymbol{A}, \gamma\,\Delta\boldsymbol{B}]\|_F.$$

For a single right-hand ($d = 1$), Rao [9] formulated a weighted TLS, and Paige and Strakoš [7] formulated a scaled TLS problem, which uses a scale factor γ to relate $\boldsymbol{A}$ and $\boldsymbol{B}$. The solution to their scaled problem is the TLS solution when $\gamma = 1$, approaches the solution to the LS problem as $\gamma \to 0$, and approaches the solution to the DLS problem as $\gamma \to \infty$. The underlying statistical assumption behind these methods is that the true error matrices for $\boldsymbol{A}$ and $\boldsymbol{B}$ are column-wise uncorrelated, and the columns of $\boldsymbol{A}$ have variance not necessarily identical to that of the columns of $\boldsymbol{B}$. In order to correctly obtain an estimate for $\boldsymbol{X}$, the covariance matrices must be known except for the single scaling constant γ that relates the two variances. However, neither [9] nor [7] discusses how to determine the scaling factor.

The main results of our work are as follows. We define in Section 2 an implicitly-weighted TLS formulation (ITLS) that reparameterizes these formulations to make computation easier. In particular, we use a scaling constant that ranges between 0 and 1 rather than the less convenient 0 and ∞. We propose in Section 3 an efficient method for computing the family of solutions. We prove asymptotic properties of the solution (as $m \to \infty$) in Section 4, holding even for rank-deficient problems. With this guidance, we propose algorithms for parameter choice in Section 5. We provide experimental results on the usefulness of ITLS in Section 6.

A simple notational convention will be helpful: a matrix $\boldsymbol{E}_C$ always denotes the true error in the matrix $\boldsymbol{C}$, and a matrix $\Delta\boldsymbol{C}$ always denotes our correction matrix for $\boldsymbol{C}$. We denote by $\widetilde{\boldsymbol{X}}$ the true parameters for our model, by $\boldsymbol{X}$ an estimated set of parameters, and by $\widehat{\boldsymbol{X}}$ a TLS estimate. Also, $\boldsymbol{I}_p$ denotes an identity matrix of dimension p.

2. Implicitly weighted total least squares

In this section, we define the ITLS problem and show its relation to previous problem formulations. Perhaps most importantly, we discuss the error assumption that makes the ITLS formulation reasonable.

2.1. ITLS and other estimation methods

Our underlying data model for ITLS is the following:

$$(\boldsymbol{A} - (1-\alpha)\boldsymbol{E}_{A_w})\widetilde{\boldsymbol{X}} = (\boldsymbol{B} - \alpha\boldsymbol{E}_{B_w}), \tag{2.1}$$

where matrices $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{B} \in \mathbb{R}^{m\times d}$ are given, α is a given weighting parameter satisfying $\alpha \in [0, 1]$, and $\boldsymbol{E}_{A_w}$ and $\boldsymbol{E}_{B_w}$ are the scaled errors in $\boldsymbol{A}$ and $\boldsymbol{B}$. We want to estimate the matrix $\widetilde{\boldsymbol{X}}$, the true values of the model's parameters.

Table 1
Relations between ITLS and STLS.

Estimation method	ITLS	STLS
Data least squares	$\alpha = 0$	$\gamma \to \infty$
Total least squares	$\alpha = 0.5$	$\gamma = 1$
Least squares	$\alpha = 1$	$\gamma \to 0$

Given that model, we define the ITLS problem as follows:

$$\min_{\boldsymbol{X}, \Delta \boldsymbol{A}_w, \Delta \boldsymbol{B}_w} \|[\Delta \boldsymbol{A}_w, \Delta \boldsymbol{B}_w]\|_F \tag{2.2}$$

subject to

$$(\boldsymbol{A} - (1-\alpha)\Delta \boldsymbol{A}_w)\boldsymbol{X} = (\boldsymbol{B} - \alpha \Delta \boldsymbol{B}_w). \tag{2.3}$$

The matrices $\Delta \boldsymbol{A}_w$ and $\Delta \boldsymbol{B}_w$ are corrections corresponding to $\boldsymbol{E}_{A_w}$ and $\boldsymbol{E}_{B_w}$. The following lemma explains how the ITLS formulation unifies DLS, LS, and TLS.

Lemma 2.1. *The ITLS defined by (2.2) and (2.3) is equivalent to DLS when $\alpha = 0$, LS when $\alpha = 1$, and TLS when $\alpha = 1/2$.*

Proof. If $\alpha = 0$, then the matrix $\Delta \boldsymbol{B}_w$ does not contribute to (2.3), so its optimal value is $\Delta \boldsymbol{B}_w = \boldsymbol{0}$, and ITLS reduces to the data least squares problem DLS. Similarly, if $\alpha = 1$, then the optimal value of $\Delta \boldsymbol{A}_w$ is $\boldsymbol{0}$ and ITLS reduces to the least squares problem LS. If $\alpha = 1/2$, then we see by defining $\Delta \boldsymbol{A} = \Delta \boldsymbol{A}_w/2$ and $\Delta \boldsymbol{B} = \Delta \boldsymbol{B}_w/2$ that the problem is equivalent to TLS, and the value of our objective function (2.2) is two times the norm of the correction term $[\Delta \boldsymbol{A}, \Delta \boldsymbol{B}]$ in (1.2). □

In the case of a single right-hand side ($d = 1$), Paige and Strakoš [7] devised a scaled TLS (STLS) formulation. We can easily extend their formulation to the case of multiple right-hand-side data: For a given $\gamma \in (0, \infty)$,

$$\min_{\boldsymbol{X}, \Delta \boldsymbol{A}_s, \Delta \boldsymbol{B}_s} \|[\Delta \boldsymbol{A}_s, \Delta \boldsymbol{B}_s]\|_F \text{ s.t. } (\boldsymbol{A} - \Delta \boldsymbol{A}_s)\boldsymbol{X}\gamma = (\boldsymbol{B}\gamma - \Delta \boldsymbol{B}_s) \tag{2.4}$$

Paige and Strakoš proved that STLS becomes LS as $\gamma \to 0$, DLS as $\gamma \to \infty$, and TLS when $\gamma = 1$. The equivalence between ITLS and STLS for these three cases is summarized in Table 1. The following lemma establishes equivalence for other values of $\gamma \in (0, \infty)$.

Lemma 2.2 (Relation between α and γ)**.** *ITLS in (2.2) and STLS in (2.4) are equivalent to each other when the parameters α and γ satisfy*

$$\gamma = \frac{1-\alpha}{\alpha} \in (0, \infty). \tag{2.5}$$

Proof. Dividing the constraint equation in (2.4) by γ, we obtain

$$(\boldsymbol{A} - \Delta \boldsymbol{A}_s)\boldsymbol{X} = \left(\boldsymbol{B} - \frac{\Delta \boldsymbol{B}_s}{\gamma}\right).$$

By defining $\Delta \boldsymbol{A}_w$ and $\Delta \boldsymbol{B}_w$ by

$$\Delta \boldsymbol{A}_s = (1-\alpha)\Delta \boldsymbol{A}_w \quad \text{and} \quad \Delta \boldsymbol{B}_s = (1-\alpha)\Delta \boldsymbol{B}_w, \tag{2.6}$$

we can rewrite the equation above as

$$(\boldsymbol{A} - (1-\alpha)\Delta \boldsymbol{A}_w)\,\boldsymbol{X} = \left(\boldsymbol{B} - \frac{(1-\alpha)\Delta \boldsymbol{B}_w}{\gamma}\right).$$

By using (2.5) in the equation above, we obtain the constraint equation (2.3). Moreover, by substituting (2.6) in the minimization in (2.4), we obtain

$$\min_{\boldsymbol{X},\Delta\boldsymbol{A}_w,\Delta\boldsymbol{B}_w} |||[(1-\alpha)\Delta\boldsymbol{A}_w, (1-\alpha)\Delta\boldsymbol{B}_w]|||_F,$$

which is equivalent to (2.2) since $(1-\alpha)$ is a fixed constant. □

Even though ITLS and STLS are mathematically equivalent, notice that the parameter α in (2.3) ranges over [0, 1] while γ in (2.4) ranges over $(0,\infty)$. A main theme in this paper is the optimal choice of parameter value. Many robust algorithms (e.g., golden section search) can be applied only to optimization problems on bounded domains, so changing the parameterization from γ to α gives a key computational advantage. For this reason, the ITLS formulation is preferable to STLS.

2.2. ITLS and the error assumption

Now we develop an error assumption consistent with the ITLS formulation and explain the statistical meaning of the weight α. This will clarify when and how ITLS can be used.

Suppose we have a model $\boldsymbol{KZ} \approx \boldsymbol{Y}$, with errors in both the model matrix $\boldsymbol{K}$ and the observations $\boldsymbol{Y}$. As before, we want to estimate the variables $\boldsymbol{Z}$ and the correction matrices $\Delta\boldsymbol{K}$ and $\Delta\boldsymbol{Y}$ satisfying

$$(\boldsymbol{K} - \Delta\boldsymbol{K})\boldsymbol{Z} = (\boldsymbol{Y} - \Delta\boldsymbol{Y}). \tag{2.7}$$

We want to formulate this as an *errors-in-variable* (EIV) problem [11, Section 8.4]. Such a formulation, from the statistical literature, is closely related to TLS but makes some extra assumptions on the errors. In particular, the rows of the error matrices should be independent, uncorrelated, and identically distributed with finite variance. Under these assumptions, if the noise-free problem has a solution, then the solution to the ITLS problem converges to the true solution with probability 1 as $m \to \infty$, as we will show in Section 4.

The independence of the error rows can be imposed by pre-multiplying (2.7) by an appropriate matrix $\boldsymbol{D} \in \mathbb{R}^{m\times m}$. We assume that this pre-multiplication has already been done, so that currently $\boldsymbol{D} = \boldsymbol{I}_m$.

To make the columns of the error uncorrelated with constant variance, we need an estimate of the covariance matrix for the errors $[\boldsymbol{E}_K, \boldsymbol{E}_Y]$. We consider the case in which the errors in $\boldsymbol{K}$ are uncorrelated with the errors in $\boldsymbol{Y}$, so the covariance matrix is block diagonal:

$$\mathrm{Cov}[\boldsymbol{E}_K, \boldsymbol{E}_Y] = \begin{bmatrix} \sigma_A^2\widehat{\boldsymbol{C}}_K & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_B^2\widehat{\boldsymbol{C}}_Y \end{bmatrix}. \tag{2.8}$$

We assume that we have good estimates of the nonsingular matrices $\widehat{\boldsymbol{C}}_K \in \mathbb{R}^{n\times n}$ and $\widehat{\boldsymbol{C}}_Y \in \mathbb{R}^{d\times d}$ but that one or both of the scalars σ_A^2 and σ_B^2 may be unknown. (Often, $\widehat{\boldsymbol{C}}_K$ and $\widehat{\boldsymbol{C}}_Y$ are estimated as identity matrices.)

Let $\widehat{\boldsymbol{C}}_K = \boldsymbol{L}_K\boldsymbol{L}_K^T$ and $\widehat{\boldsymbol{C}}_Y = \boldsymbol{L}_Y\boldsymbol{L}_Y^T$, where $\boldsymbol{L}_K$ and $\boldsymbol{L}_Y$ are Cholesky factors. Define

$$\begin{aligned}
\boldsymbol{A} &= \boldsymbol{K}\boldsymbol{L}_K^{-T}, \\
\boldsymbol{B} &= \boldsymbol{Y}\boldsymbol{L}_Y^{-T}, \\
\boldsymbol{E}_A &= \boldsymbol{E}_K\boldsymbol{L}_K^{-T}, \\
\boldsymbol{E}_B &= \boldsymbol{E}_Y\boldsymbol{L}_Y^{-T}, \\
\boldsymbol{X} &= \boldsymbol{L}_K^T\boldsymbol{Z}\boldsymbol{L}_Y^{-T}.
\end{aligned}$$

Under these definitions, it is easy to verify that the constraint (2.7) is equivalent to the constraint (1.1) studied above. By the construction of $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$, the covariance matrix for the transformed errors becomes

$$\mathrm{Cov}[\boldsymbol{E}_A, \boldsymbol{E}_B] = \begin{bmatrix} \sigma_A^2\boldsymbol{I}_n & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_B^2\boldsymbol{I}_d \end{bmatrix}.$$

To satisfy the assumptions in [4] for EIV convergence, we need only scale so that the variances are identical. To do this, we define

$$\sigma_E = \sigma_A + \sigma_B,$$
$$\alpha = \sigma_B/\sigma_E.$$

Then $0 < \alpha < 1$ (as long as both σ_A^2 and σ_B^2 are positive), and $1 - \alpha = \sigma_A/\sigma_E$. Now let

$$\boldsymbol{A}_\alpha = \alpha \boldsymbol{A},$$
$$\boldsymbol{B}_\alpha = (1-\alpha)\boldsymbol{B}.$$

Then the corresponding (true) errors $\boldsymbol{E}_{A_\alpha} = \alpha \boldsymbol{E}_A$ and $\boldsymbol{E}_{B_\alpha} = (1-\alpha)\boldsymbol{E}_B$ are uncorrelated and have identical variances $\sigma_A^2\sigma_B^2/\sigma_E^2$. Finally, we obtain a linear model containing uncorrelated errors with identical variances:

$$(\boldsymbol{A}_\alpha - \boldsymbol{E}_{A_\alpha})\boldsymbol{X}_\alpha = (\boldsymbol{B}_\alpha - \boldsymbol{E}_{B_\alpha}), \tag{2.9}$$

where

$$\boldsymbol{E}_{A_\alpha} = \frac{\sigma_B}{\sigma_E}\boldsymbol{E}_A, \quad \boldsymbol{E}_{B_\alpha} = \frac{\sigma_A}{\sigma_E}\boldsymbol{E}_B, \quad \text{and} \quad \boldsymbol{X}_\alpha = \left(\frac{\sigma_A}{\sigma_B}\right)\boldsymbol{X}. \tag{2.10}$$

The matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ can be determined from the observed data matrices $(\boldsymbol{K}, \boldsymbol{Y})$ and the Cholesky factors $(\boldsymbol{L}_K, \boldsymbol{L}_Y)$, but $\boldsymbol{A}_\alpha$ and $\boldsymbol{B}_\alpha$ contain the parameters σ_A^2 and σ_B^2.

Using the linear model (2.9), we can formulate a TLS problem, which includes the ratio σ_A^2/σ_B^2:

$$\min_{\boldsymbol{X}_\alpha, \Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha} \|\Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha\|_F \text{ s.t. } (\boldsymbol{A}_\alpha - \Delta\boldsymbol{A}_\alpha)\boldsymbol{X}_\alpha = (\boldsymbol{B}_\alpha - \Delta\boldsymbol{B}_\alpha). \tag{2.11}$$

We have thus proven the following lemma.

Lemma 2.3 (ITLS and equivalent TLS). *If $\sigma_A^2 > 0$ and $\sigma_B^2 > 0$, then the TLS problem (2.11) is equivalent to ITLS (2.2) and (2.3) when $\alpha \in (0,1)$ satisfies*

$$\frac{\sigma_A}{\sigma_B} = \frac{1-\alpha}{\alpha}. \tag{2.12}$$

Paige and Strakoš [7] also made use of σ_A^2/σ_B^2 in defining γ for their STLS formulation.

We see that if we know the ratio of σ_A^2 to σ_B^2, then we can estimate the desired solution by solving the ITLS problem with $\alpha = \sigma_B/\sigma_E$. If $\sigma_A^2 = \sigma_B^2$, then $\alpha = 1/2$ and we have the standard TLS problem. For small values of the ratio, $\alpha \approx 1$ and we solve a problem close to LS. For large values, $\alpha \approx 0$ and we solve a problem close to DLS.

If the ratio σ_A^2/σ_B^2 is not known, then it is not clear what value of α should be used. We propose an answer to this dilemma in Section 5, using a method that varies α. In order to make this practical, we need an efficient algorithm for solving ITLS for multiple values of α. We develop such an algorithm in the next section.

3. Computing ITLS solutions

In this section, we show that after an initial decomposition of the $m \times (n+d)$ matrix $[\boldsymbol{A}_\alpha, \boldsymbol{B}_\alpha]$, we can compute the solution to the ITLS problem for any other value of α by working with a smaller upper-triangular matrix of dimension $(n+d) \times (n+d)$ when $m > n+d$.

3.1. Reduction of the problem

Following well-known results for the standard TLS problem, as described in [11, Chapter 2 and 3], we begin with some notation. Define the SVD of

$$[\boldsymbol{A}_\alpha, \boldsymbol{B}_\alpha] = [\alpha\boldsymbol{A}, (1-\alpha)\boldsymbol{B}] \in \mathbb{R}^{m\times(n+d)}$$

by

$$[\boldsymbol{A}_\alpha, \boldsymbol{B}_\alpha] = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T = \underset{t}{[\boldsymbol{U}_1}\ \underset{q}{\boldsymbol{U}_2}] \begin{bmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{\Sigma}_2 \end{bmatrix} \begin{matrix} t \\ q \end{matrix} \begin{bmatrix} \boldsymbol{V}_1^T \\ \boldsymbol{V}_2^T \end{bmatrix} \begin{matrix} t \\ q, \end{matrix} \tag{3.1}$$
$$\underset{t}{} \quad \underset{q}{\phantom{\boldsymbol{U}_2]}} \qquad \underset{t}{\phantom{\boldsymbol{\Sigma}_1}} \quad \underset{q}{\phantom{\boldsymbol{\Sigma}_2}} \qquad\qquad \underset{n+d}{\phantom{\boldsymbol{V}_1^T}}$$

where $\boldsymbol{U} = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_{n+d}] \in \mathbb{R}^{m\times(n+d)}$ and $\boldsymbol{V} = [\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n+d}] \in \mathbb{R}^{(n+d)\times(n+d)}$ have orthonormal columns, $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_{n+d})$, $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_{n+d} \geqslant 0$, t is an integer in $[0, n+d]$, and $t + q = n + d$.

Let $\widehat{\boldsymbol{A}}_\alpha$ and $\widehat{\boldsymbol{B}}_\alpha$ denote the corrected matrices

$$\widehat{\boldsymbol{A}}_\alpha = \boldsymbol{A}_\alpha - \Delta\boldsymbol{A}_\alpha \quad \text{and} \quad \widehat{\boldsymbol{B}}_\alpha = \boldsymbol{B}_\alpha - \Delta\boldsymbol{B}_\alpha, \tag{3.2}$$

for some correction matrices $\Delta\boldsymbol{A}_\alpha$ and $\Delta\boldsymbol{B}_\alpha$. Define $\widehat{\boldsymbol{X}}_\alpha$ to be the TLS solution (if it exists) associated with the corrected matrices $\widehat{\boldsymbol{A}}_\alpha$ and $\widehat{\boldsymbol{B}}_\alpha$, satisfying

$$\widehat{\boldsymbol{A}}_\alpha\widehat{\boldsymbol{X}}_\alpha = \widehat{\boldsymbol{B}}_\alpha \quad \text{or} \quad [\widehat{\boldsymbol{A}}_\alpha, \widehat{\boldsymbol{B}}_\alpha]\begin{bmatrix} \widehat{\boldsymbol{X}}_\alpha \\ -\boldsymbol{I}_d \end{bmatrix} = \boldsymbol{0}. \tag{3.3}$$

By the Eckart–Young–Mirsky Theorem, the solution to the problem

$$\min_{\mathrm{rank}([\widehat{\boldsymbol{A}}_\alpha, \widehat{\boldsymbol{B}}_\alpha])=t} \|[\Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha]\|_F^2 \tag{3.4}$$

is

$$[\widehat{\boldsymbol{A}}_\alpha, \widehat{\boldsymbol{B}}_\alpha] = \boldsymbol{U}_1\boldsymbol{\Sigma}_1\boldsymbol{V}_1^T, \tag{3.5}$$

and the value of the minimization function is

$$\sum_{i=t+1}^{n+d} \sigma_i^2.$$

The corresponding correction matrix $[\Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha]$ is

$$[\Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha] = \boldsymbol{U}_2\boldsymbol{\Sigma}_2\boldsymbol{V}_2^T. \tag{3.6}$$

Because of this, the solution $\widehat{\boldsymbol{X}}_\alpha$ of (3.3) must satisfy

$$\mathrm{Range}\left(\begin{bmatrix} \widehat{\boldsymbol{X}}_\alpha \\ -\boldsymbol{I}_d \end{bmatrix}\right) \subseteq \mathrm{Null}\,(\boldsymbol{V}_1^T) = \mathrm{Range}(\boldsymbol{V}_2), \tag{3.7}$$

by orthogonality of the right singular matrix $\boldsymbol{V}$. In order to determine an appropriate partition size t, we need to consider (i) the existence of $\widehat{\boldsymbol{X}}_\alpha$ and (ii) the noise level. We partition $\boldsymbol{V}_2$ as

$$\boldsymbol{V}_2 = \begin{matrix} \\ \\ \end{matrix} \begin{bmatrix} \boldsymbol{V}_{12} \\ \boldsymbol{V}_{22} \end{bmatrix} \begin{matrix} n \\ d. \end{matrix} \tag{3.8}$$
$$\underset{q}{\phantom{\begin{bmatrix} \boldsymbol{V}_{12} \\ \boldsymbol{V}_{22} \end{bmatrix}}}$$

Further, let $\hat{t}$ denote our choice of t and $\hat{q}$ denote the corresponding q, so that $\hat{t} + \hat{q} = n + d$.

First, for a given t, such a $\widehat{\boldsymbol{X}}_\alpha$ may not exist unless the block matrix $\boldsymbol{V}_{22}$ of the last d rows in the corresponding matrix $\boldsymbol{V}_2$, has column rank d. Therefore we want

$$\hat{t} \leqslant t_0 \quad \text{where} \quad t_0 = \max\{t : \mathrm{rank}(\boldsymbol{V}_{22}) = d\}. \tag{3.9}$$

Second, we would like the magnitude of the correction term to be less than a given noise tolerance ϵ:

$$\|\Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha\|_F^2 = \sum_{i=t+1}^{n+d} \sigma_i^2 < \epsilon. \tag{3.10}$$

Let r be the minimal value of t satisfying the inequality above, which is called the *numerical rank*. Then we choose

$$\hat{t} = \min(t_0, r). \tag{3.11}$$

Note that, if such $\hat{t}$ is less than n, there exist infinitely many solutions $\widehat{\boldsymbol{X}}_\alpha$ satisfying (3.3) or (3.7). In this case, we can single out a minimal norm solution among these candidates.

Let $\widetilde{\boldsymbol{V}}_2 \in \mathbb{R}^{(n+d)\times q}$ denote a matrix containing an orthonormal basis for Range($\boldsymbol{V}_2$), and partition $\widetilde{\boldsymbol{V}}_2$ as

$$\widetilde{\boldsymbol{V}}_2 = \begin{matrix}\begin{bmatrix}\widetilde{\boldsymbol{V}}_{12}\\ \widetilde{\boldsymbol{V}}_{22}\end{bmatrix} & \begin{matrix}n\\ d.\end{matrix}\\ q\end{matrix}$$

For a chosen partition size $\hat{t}$ and $\hat{q}$, we can compute a minimal norm solution $\widehat{\boldsymbol{X}}_\alpha$ and the correction term $[\Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha]$ as

$$\widehat{\boldsymbol{X}}_\alpha = -\widetilde{\boldsymbol{V}}_{12}\widetilde{\boldsymbol{V}}_{22}^{\dagger} \tag{3.12}$$

$$[\Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha] = [\boldsymbol{A}_\alpha, \boldsymbol{B}_\alpha]\widetilde{\boldsymbol{V}}_2\widetilde{\boldsymbol{V}}_2^T. \tag{3.13}$$

Thus, we can compute the minimal norm TLS solution $\widehat{\boldsymbol{X}}_\alpha$ and the corresponding correction matrix $[\Delta\boldsymbol{A}_\alpha, \Delta\boldsymbol{B}_\alpha]$ solely from $\widetilde{\boldsymbol{V}}_2$, a matrix whose column space is the partial right singular subspace, without necessarily computing the right singular matrix $\boldsymbol{V}_2$.

3.2. Economical computation of $\widetilde{\boldsymbol{V}}_2$

We now consider how the basis matrix $\widetilde{\boldsymbol{V}}_2$ can be computed. Clearly we could use the standard Golub–Kahan algorithm [5] to compute the SVD of $[\boldsymbol{A}_\alpha, \boldsymbol{B}_\alpha]$, obtaining the basis $\widetilde{\boldsymbol{V}}_2 = \boldsymbol{V}_2$, but there are more economical alternatives when multiple values of α are of interest. For example, the rank-revealing ULV algorithm [10] can accurately compute this basis without producing the SVD, and it was used in [3] to solve the TLS problem. Other alternatives include the partial SVD method (PSVD) [11, Section 4.3] and the implicitly-restarted Arnoldi algorithm [6].

If $m > n + d$, it is desirable to apply one of these algorithms to a smaller matrix. For example, we could first compute the $(n + d) \times (n + d)$ upper-triangular factor $\boldsymbol{R}_\alpha$ from the QR decomposition of $[\boldsymbol{A}_\alpha, \boldsymbol{B}_\alpha]$. According to [2], using QR before SVD reduces the computational cost when $m > \frac{5}{3}(n + d)$.

While searching for an appropriate value of α for ITLS, we need to compute the SVD of $[\boldsymbol{A}_\alpha, \boldsymbol{B}_\alpha]$ for different values of α. For a new parameter value α', the new upper-triangular factor is

$$\boldsymbol{R}'_\alpha = \boldsymbol{R}_\alpha \begin{bmatrix} \left(\frac{\alpha'}{\alpha}\right)\boldsymbol{I}_n & \boldsymbol{0} \\ \boldsymbol{0} & \left(\frac{1-\alpha'}{1-\alpha}\right)\boldsymbol{I}_d \end{bmatrix}. \tag{3.14}$$

The cost of this scaling is only $O((n + d)^2)$, rather than the $O(m(n + d)^2)$ cost needed to compute the QR decomposition of $[\boldsymbol{A}_{\alpha'}, \boldsymbol{B}_{\alpha'}]$. Thus, we will compute the right singular subspace of $\boldsymbol{R}_\alpha$ instead of $[\boldsymbol{A}_\alpha, \boldsymbol{B}_\alpha]$ for different weights α.

In Section 5 we propose a method for choosing an optimal value of α, and this requires computing $\widetilde{\boldsymbol{V}}_2$ for many candidate values of α. In such an algorithm, it is especially important to economize by using (3.14) in conjunction with an algorithm such as the PSVD.

4. Asymptotic behavior

In this section we keep α fixed but let m, the number of observations, vary, so our notation will change to reflect this. We study the behavior of the ITLS problem as $m \to \infty$. Our development follows that of Gleser[1] [4] except that *we also treat the rank-deficient case*.

[1] Gleser's X^T, U^T, and B^T correspond to our $[\boldsymbol{A}, \boldsymbol{B}]$, $[\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}]$, and $\widetilde{\boldsymbol{X}}$ respectively, and we set his α to zero.

Let $[\widetilde{\boldsymbol{A}}_m, \widetilde{\boldsymbol{B}}_m]$ denote the true but unknown matrix, and suppose it has rank $r \leqslant n$. (Since the columns of $\widetilde{\boldsymbol{B}}_m$ are in the range of $\widetilde{\boldsymbol{A}}_m$, the rank cannot be greater than n.) Let $\widetilde{\boldsymbol{X}}$ denote the unique true solution if $r = n$, or the unique minimum norm true solution otherwise, so that

$$\widetilde{\boldsymbol{A}}_m \widetilde{\boldsymbol{X}} = \widetilde{\boldsymbol{B}}_m. \tag{4.1}$$

Then the observed data satisfy

$$[\boldsymbol{A}_m, \boldsymbol{B}_m] = [\widetilde{\boldsymbol{A}}_m, \widetilde{\boldsymbol{B}}_m] + [\boldsymbol{E}_{A,m}, \boldsymbol{E}_{B,m}] = \widetilde{\boldsymbol{A}}_m[\boldsymbol{I}_n, \widetilde{\boldsymbol{X}}] + [\boldsymbol{E}_{A,m}, \boldsymbol{E}_{B,m}].$$

Now the matrix $\widetilde{\boldsymbol{A}}_m[\boldsymbol{I}_n, \widetilde{\boldsymbol{X}}]$ also has rank r, so $[\boldsymbol{A}_m, \boldsymbol{B}_m]$ should have $(n + d - r)$ small singular values, resulting from the perturbations $[\boldsymbol{E}_{A,m}, \boldsymbol{E}_{B,m}]$. We need some insight into the behavior of these singular values.

We impose two assumptions.

Assumption 1. Each row of $[\boldsymbol{E}_{A,m}, \boldsymbol{E}_{B,m}]$ is independent and identically distributed, with zero means and covariance matrix $\sigma_\epsilon^2 \boldsymbol{I}_{n+d}$.

Assumption 2. The matrices $(1/m)\, \widetilde{\boldsymbol{A}}_m^T \widetilde{\boldsymbol{A}}_m$ converge to a finite limit Δ:

$$\lim_{m\to\infty} \frac{1}{m} \widetilde{\boldsymbol{A}}_m^T \widetilde{\boldsymbol{A}}_m = \Delta. \tag{4.2}$$

We define

$$\boldsymbol{W}_m = [\boldsymbol{A}_m, \boldsymbol{B}_m]^T [\boldsymbol{A}_m, \boldsymbol{B}_m], \tag{4.3}$$

$$\widetilde{\boldsymbol{W}}_m = [\widetilde{\boldsymbol{A}}_m, \widetilde{\boldsymbol{B}}_m]^T [\widetilde{\boldsymbol{A}}_m, \widetilde{\boldsymbol{B}}_m] = \begin{bmatrix} \boldsymbol{I}_n \\ \widetilde{\boldsymbol{X}}^T \end{bmatrix} \widetilde{\boldsymbol{A}}_m^T \widetilde{\boldsymbol{A}}_m [\boldsymbol{I}_n, \widetilde{\boldsymbol{X}}], \tag{4.4}$$

and study the convergence of these matrices.

Lemma 4.1. *Under Assumptions 1 and 2, both* $(1/m)\widetilde{\boldsymbol{W}}_m$ *and* $(1/m)\boldsymbol{W}_m$ *converge to limits*:[2]

$$\lim_{m\to\infty} \frac{1}{m} \widetilde{\boldsymbol{W}}_m = \begin{bmatrix} \boldsymbol{I}_n \\ \widetilde{\boldsymbol{X}}^T \end{bmatrix} \Delta [\boldsymbol{I}_n, \widetilde{\boldsymbol{X}}] \equiv \widetilde{\boldsymbol{\Theta}}, \tag{4.5}$$

$$\operatorname*{plim}_{m\to\infty} \frac{1}{m} \boldsymbol{W}_m = \sigma_\epsilon^2 \boldsymbol{I}_{n+d} + \widetilde{\boldsymbol{\Theta}} \equiv \boldsymbol{\Theta}. \tag{4.6}$$

Proof. The first result follows from using (4.2) in (4.5). For the second, see [4, Lemma 3.1]. □

Next, we need an eigendecomposition of $\boldsymbol{\Theta}$ and its relation to that of $\Delta(\boldsymbol{I}_n + \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^T)$.

Lemma 4.2. *Denote the eigenvalues of* $\Delta(\boldsymbol{I}_n + \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^T)$ *by* $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \geqslant 0$, *and let the columns of* $\boldsymbol{\psi}$ *be the corresponding eigenvectors. Then we have an eigendecomposition of* $\boldsymbol{\Theta}$ *as*

$$\boldsymbol{\Theta}[\boldsymbol{V}_{\Theta_1}, \boldsymbol{V}_{\Theta_2}] = [\boldsymbol{V}_{\Theta_1}, \boldsymbol{V}_{\Theta_2}] \begin{bmatrix} \sigma_\epsilon^2 \boldsymbol{I}_n + \boldsymbol{D}_\lambda & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_\epsilon^2 \boldsymbol{I}_d \end{bmatrix},$$

where $\boldsymbol{D}_\lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ *and the columns of* $\boldsymbol{V}_\Theta = [\boldsymbol{V}_{\Theta_1}, \boldsymbol{V}_{\Theta_2}]$, *with*

$$\boldsymbol{V}_{\Theta_1} = \begin{bmatrix} \boldsymbol{I}_n \\ \widetilde{\boldsymbol{X}}^T \end{bmatrix} \boldsymbol{\psi}, \quad \boldsymbol{V}_{\Theta_2} = \begin{bmatrix} -\widetilde{\boldsymbol{X}} \\ \boldsymbol{I}_d \end{bmatrix} (\boldsymbol{I}_d + \widetilde{\boldsymbol{X}}^T \widetilde{\boldsymbol{X}})^{-\frac{1}{2}}, \tag{4.7}$$

are mutually orthogonal and have norm 1.

[2] We denote "convergence with probability one" using the notation "plim".

Proof (*See [4, p. 35]*). The symmetric positive semi-definite matrix $(\boldsymbol{I}_n + \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^T)^{\frac{1}{2}}\Delta(\boldsymbol{I}_n + \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^T)^{\frac{1}{2}}$ has eigenvalues that are real and non-negative and has an eigenvector matrix, denoted by $(\boldsymbol{I}_n + \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^T)^{\frac{1}{2}}\boldsymbol{\psi}$, that is orthonormal:

$$\boldsymbol{\psi}^T(\boldsymbol{I}_n + \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^T)\boldsymbol{\psi} = \boldsymbol{I}_n. \tag{4.8}$$

The matrix $\Delta(\boldsymbol{I}_n + \widetilde{\boldsymbol{X}}\widetilde{\boldsymbol{X}}^T)$ is similar to this matrix and has eigenvectors $\boldsymbol{\psi}$.

The eigendecomposition of $\boldsymbol{\Theta}$ and the orthonormality of its eigenbasis are verified by direct computation. □

Using this eigendecomposition, we can understand the convergence of the singular values σ_i from (3.1).

Lemma 4.3. *Let $\sigma_{1,m} \geqslant \sigma_{2,m} \geqslant \cdots \geqslant \sigma_{n+d,m} \geqslant 0$ denote the singular values of $[\boldsymbol{A}_m, \boldsymbol{B}_m]$. Under Assumptions 1 and 2,*

$$\operatorname*{plim}_{m\to\infty} \frac{1}{m}\sigma_{i,m}^2 = \begin{cases} \sigma_\epsilon^2 + \lambda_i, & i = 1,\ldots,n, \\ \sigma_\epsilon^2, & i = n+1,\ldots,n+d. \end{cases}$$

Proof. This is a direct consequence of the definition of $\boldsymbol{W}_m$ in (4.3), the convergence of $(1/m)\boldsymbol{W}_m$ to $\boldsymbol{\Theta}$ (Lemma 4.1), and Lemma 4.2. □

Gleser [4, Assumption C] assumes that Δ is positive definite, but we are able to omit that assumption. We denote the rank of the symmetric positive semi-definite matrix Δ by $r \leqslant n$. Then $\lambda_i = 0$ for $i = r+1,\ldots,n$, so by Lemma 4.3,

$$\operatorname*{plim}_{m\to\infty} \frac{1}{m}\sigma_{i,m}^2 = \sigma_\epsilon^2 \quad \text{for} \quad i = r+1,\ldots,n+d. \tag{4.9}$$

This gives us a way to estimate σ_ϵ^2, as shown in the following lemma.

Lemma 4.4. *Let*

$$\widehat{\sigma}_{\epsilon,m}^2 = \frac{1}{n+d-r}\sum_{i=r+1}^{n+d}\sigma_{i,m}^2. \tag{4.10}$$

Under Assumptions 1 and 2,

$$\operatorname*{plim}_{m\to\infty} \frac{1}{m}\widehat{\sigma}_{\epsilon,m}^2 = \sigma_\epsilon^2.$$

Proof. This is a direct result of Lemma 4.3 and the fact that $\lambda_i = 0$ for $i = r+1,\ldots,n$. □

In order to use $\widehat{\sigma}_{\epsilon,m}^2$ in an algorithm, we need to know that we can reliably estimate the rank r as $m \to \infty$.

Lemma 4.5. *Under Assumptions 1 and 2,*

$$\lim_{m\to\infty} \Pr\left\{\frac{1}{m}(\sigma_{r,m}^2 - \sigma_{r+1,m}^2) < \frac{\lambda_r}{2}\right\} = 0.$$

Proof. The result follows since $(1/m)(\sigma_{r,m}^2 - \sigma_{r+1,m}^2)$ converges with probability one to $\lambda_r > 0$. □

With this result and (4.10), we see that, with appropriate choice of ϵ in (3.10), our rank estimation algorithm in (3.11) gives the correct result (with probability one) as $m \to \infty$, and from

this we can establish convergence of the solution estimates, just as Gleser did in the full-rank case [4, Lemma 3.3].

Lemma 4.6. *Under Assumptions 1 and 2,*

$$\operatorname*{plim}_{m\to\infty} \widehat{\boldsymbol{X}}_m = \widetilde{\boldsymbol{X}}$$

where $\widetilde{\boldsymbol{X}}$ is the minimal norm true solution satisfying (4.1) and ϵ in (3.10) satisfies

$$m(n+d-r)\sigma_\epsilon^2 \leqslant \epsilon \leqslant m\left((n+d-r+1)\sigma_\epsilon^2 + \frac{\lambda_r}{2}\right). \tag{4.11}$$

Proof. With this choice of ϵ, by Lemma 4.5, our estimated rank converges to the true rank r with probability one. Since $(1/m)\boldsymbol{W}_m$ converges with probability one to $\boldsymbol{\Theta}$, and since there is, by Lemma 4.2, a gap in the spectrum of $\boldsymbol{\Theta}$, the invariant subspace corresponding to the smallest $n+d-r$ eigenvalues of $(1/m)\boldsymbol{W}_m$ converges with probability one to the span of the last $n+d-r$ columns of $\boldsymbol{V}_\Theta$. Since our estimate $\widehat{\boldsymbol{X}}_m$ is independent of the choice of basis for this invariant subspace, it also must converge with probability one to $\widetilde{\boldsymbol{X}}$, which, by (4.7) and the formula (3.12), is the desired minimum norm solution. □

We have now laid the groundwork for algorithms for choosing ITLS parameters. From Lemma 4.1, we know that the sequence of $\boldsymbol{W}$ matrices converges with probability one to $\boldsymbol{\Theta}$, and from (4.7) we know that $\boldsymbol{V}_{\Theta_2}$ is full rank. Therefore, our parameter t_0 in (3.9) converges with probability one to n, so $\hat{t}$ in (3.11) converges to r. From now on, we assume, based on Lemma 4.4 and Lemma 4.5, that we have enough observations so that in (3.11) we have $\hat{t} = r$, when ϵ in (3.10) satisfies (4.11).

5. Choice of parameters

In this section, we propose two heuristic methods to determine the ITLS parameters based on the asymptotic convergence properties established in the previous section. We consider two cases: (1) either σ_A^2 or σ_B^2 is known, or (2) neither is known, in which case we require $n+d-r > 1$.

5.1. Prior information on σ_A^2 or σ_B^2

If the weight parameter α perfectly adjusts the variance of $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$, then $\boldsymbol{E}_{A_\alpha}$ and $\boldsymbol{E}_{B_\alpha}$ have identical variances, so that

$$\alpha^2\sigma_A^2 = (1-\alpha)^2\sigma_B^2 = \sigma_\epsilon^2. \tag{5.1}$$

By Lemma 4.4, $\widehat{\sigma}_\epsilon^2 = \widehat{\sigma}_{\epsilon,m}^2/m$ is a consistent estimate for σ_ϵ^2. Therefore, if we know σ_A^2, for example, then it is reasonable to find the α that minimizes a relative gap between $\alpha^2\sigma_A^2$ and $\widehat{\sigma}_\epsilon^2$:

$$\min_\alpha \left|\log \frac{\widehat{\sigma}_\epsilon^2}{\alpha^2\sigma_A^2}\right|. \tag{5.2}$$

Similarly, if we know σ_B^2, we could choose the value of α that solves the problem

$$\min_\alpha \left|\log \frac{\widehat{\sigma}_\epsilon^2}{(1-\alpha)^2\sigma_B^2}\right|. \tag{5.3}$$

Fig. 1 illustrates how the estimated error variance $\widehat{\sigma}_\epsilon^2$ changes with α. The red and blue dashed lines represent the change of $\alpha^2\sigma_A^2$ and $(1-\alpha)^2\sigma_B^2$, and their intersection gives the true α and the true σ_ϵ^2, by (5.1). We can see that the estimate $\widehat{\sigma}_\epsilon^2$ approaches the true error variance σ_ϵ^2 as α approaches the true value, illustrating the usefulness of a choice of α based on the minimization problem (5.2) or (5.3).

In order to compute $\widehat{\sigma}_\epsilon^2$, the rank r of Δ is required. If the rank is given to us, we can immediately apply the optimization methods above. If not, we also need to estimate the rank. We examine how $\widehat{\sigma}_\epsilon^2$

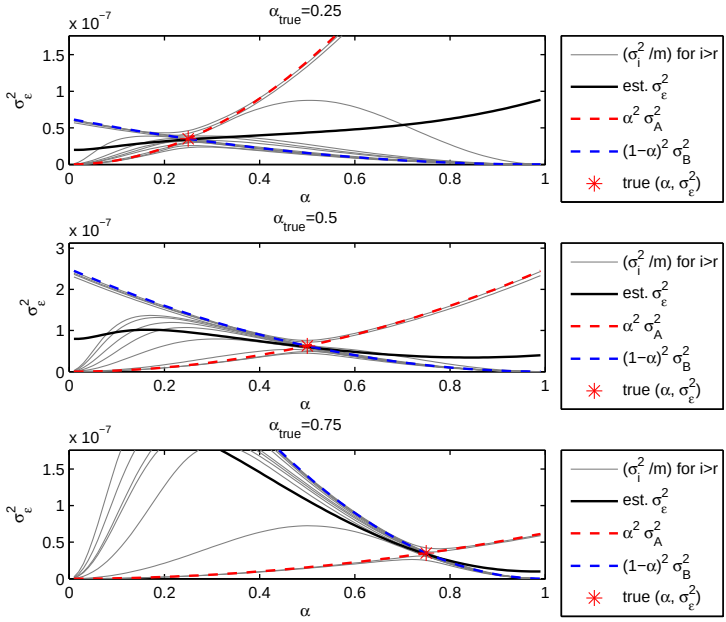


Fig. 1. The estimated error variance $\widehat{\sigma}_\epsilon^2$ as a function of α, for $\alpha_{true} = 0.25$, 0.5, and 0.75. The true value of $(\alpha, \sigma_\epsilon^2)$ is the intersection of the $\alpha^2\sigma_A^2$ curve (red dashed) and the $(1-\alpha)^2\sigma_B^2$ curve (blue dashed), marked with a star. The behavior of the small singular values as a function of α is traced by the grayish curves. The test problem is specified in Section 6, with $m = 200$, $n = 8$, $r = 6$, $d = 10$, $\sigma_E = 0.01$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

and the resulting objective function values are influenced by the estimate $\hat{r}$ of the rank. First, when $\hat{r}$ is overestimated, we expect that the minimum value of (5.2) is still close to 0. This is because $\widehat{\sigma}_\epsilon^2$ is still a consistent estimator of $\alpha^2\sigma_A^2$ when α is well estimated, as shown in (4.9). Second, if $\hat{r}$ is underestimated, the resulting $\widehat{\sigma}_\epsilon^2$ is overwhelmed by incorrectly adding large σ_i^2 for $i < r$. From these observations, we can determine $\hat{r}$ by solving (5.2), decreasing $\hat{r}$ from n to 1. We recognize the correct rank by looking for a jump and then a plateau in the optimal objective function value. A similar argument holds for (5.3). In contrast to (5.2), though, the denominator will force the minimizer α to be lower than its true value, so looking for a jump in the optimal α as $\hat{r}$ is changed is an alternative way to recognize an underestimated rank. We will exhibit these phenomena with sample problems in Section 6.

5.2. No prior information on σ_A^2 or σ_B^2

If we do not have any prior information about error variances σ_A^2 or σ_B^2, we cannot use (5.2) or (5.3). Instead, we use the convergence property (4.9) to evaluate a given α. Since all $(n + d - r)$ smallest singular values converge to a single constant value as the number of observations increases, we choose α to minimize their dispersion. Note that this convergence property holds only when Assumption 1 applies to our problem, which will be satisfied by the correct value of α. As an example, the grayish curves in Fig. 1 show how the smallest singular values change as α varies. We can see that the singular values get closer to each other near α_{true}.

We measure the dispersion using the coefficient of variation c_v, defined as

$$c_v(\boldsymbol{y}) = \frac{\text{std}(\boldsymbol{y})}{\text{mean}(\boldsymbol{y})},$$

where mean($\boldsymbol{y}$) and std($\boldsymbol{y}$) denote the mean and standard deviation of the data vector $\boldsymbol{y}$. Thus, we choose α as the solution to

$$\min_{\alpha} \; c_v\left(\left[\sigma_{r+1}^2(\alpha), \ldots, \sigma_{n+d}^2(\alpha)\right]\right). \tag{5.4}$$

There are other dispersion measures, such as standard deviation or variance. However, as the estimated α decreases to 0, the smallest singular values approach zero regardless of the true α, so these dispersion measures can be misleading. The coefficient of variation is dimensionless and therefore not subject to this limitation.

As in the optimization methods of Section 5.1, estimating c_v in (5.4) requires knowledge of the rank r. If the rank is not available, we can apply a similar rank-estimation strategy. For the true α, when $\hat{r}$ is an overestimate, c_v remains acceptably small by (4.9). On the other hand, when $\hat{r}$ is an underestimate, c_v grows significantly. Therefore, if we repeatedly solve (5.4) decreasing $\hat{r}$ from n, we can find an appropriate $\hat{r}$ by recognizing a jump in the corresponding value of c_v. In contrast to the rank and α estimation method of Section 5.1, this method requires $n + d - r > 1$, since we need at least two singular values to compute the coefficient of variation. Thus, we cannot use this method for a full-rank, single right-hand side TLS problem ($d = 1$ and $r = n$).

6. Experiments

We now present the results of some simple experiments exploring whether ITLS can be useful in data fitting problems. Since the "correct" choice of α depends on the error distributions for $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$, our questions are these:

- How sensitive is the solution $\boldsymbol{X}$ to the ITLS problem as α varies?
- Can the "correct" value of α be determined computationally?

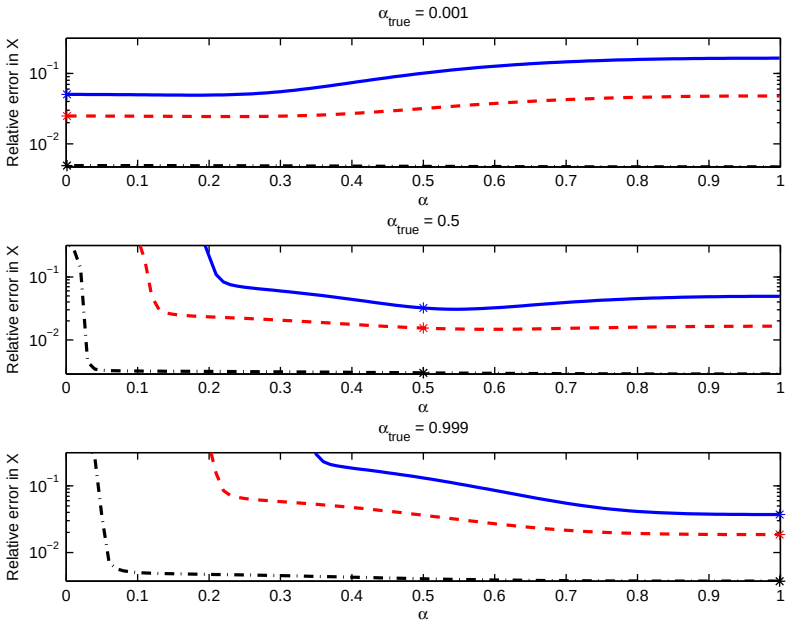


Fig. 2. Relative errors in $\boldsymbol{X}$ as a function of α for $\alpha_{true} = 0.001$, 0.5, and 0.999 with different noise levels σ_E: 0.01 (blue solid), 0.005 (red dashed), and 0.001 (black dash-dotted). The star on each curve marks α_{true}. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For given weight parameter α, rank r, and noise level σ_E, we generate a sample problem in the following way:

1. Generate $\widetilde{\boldsymbol{A}}$ and $\boldsymbol{X}$ using Matlab's `randn()`.
2. Modify $\widetilde{\boldsymbol{A}}$ to have rank r.
3. Generate $\widetilde{\boldsymbol{B}}$ as $\widetilde{\boldsymbol{B}} = \widetilde{\boldsymbol{A}}\boldsymbol{X}$.
4. Compute a minimal norm solution $\widetilde{\boldsymbol{X}}$ of $\widetilde{\boldsymbol{A}}\widetilde{\boldsymbol{X}} = \widetilde{\boldsymbol{B}}$.
5. Generate noise $\boldsymbol{E}_A \sim N(\boldsymbol{0}, \sigma_E^2(1-\alpha)^2\boldsymbol{I}_n)$ and $\boldsymbol{E}_B \sim N(\boldsymbol{0}, \sigma_E^2\alpha^2\boldsymbol{I}_d)$.[3]
6. Add the noise to $\widetilde{\boldsymbol{A}}$ and $\widetilde{\boldsymbol{B}}$ to form $\boldsymbol{A} = \widetilde{\boldsymbol{A}} + \boldsymbol{E}_A$ and $\boldsymbol{B} = \widetilde{\boldsymbol{B}} + \boldsymbol{E}_B$.

Now, given $\boldsymbol{A}$ and $\boldsymbol{B}$, our goal is to estimate the hidden parameters (α, r, σ_E) as well as the true TLS solution $\widetilde{\boldsymbol{X}}$. Note that the noise level σ_E is related to the noise variance of σ_ϵ^2 in Assumption 1 by

$$\sigma_\epsilon = \alpha(1-\alpha)\sigma_E.$$

In our first experiment, we set $m = 200$, $n = 8$, $d = 4$, $r = 6$, and varied the noise level σ_E as 0.01, 0.005, and 0.001. We obtained similar results for other choices of the problem, including non-random matrices.

First, we examine the sensitivity of the TLS solution to the choice of α. Fig. 2 plots the relative error in $\boldsymbol{X}$ as a function of α, for three different true values $\alpha_{true} = 0.001, 0.5,$ and 0.999 (which are marked by a star on the curve) with varying noise level. We can see that the sensitivity increases as the noise level increases, so the more noise, the more important it is to determine α correctly.

Next, we evaluate the performance of our methods for determining α. We apply the methods described in Section 5 to find a minimizer α for (5.2), (5.3), and (5.4), using Matlab's `fminbnd` [1], performing function evaluations using the partial SVD. The results are shown in Figs. 3(a)–5(b).

Fig. 3(a) shows the results of estimating α when σ_A is known, using the minimizer of (5.2) for different values of $\hat{r}$ with $\sigma_E = 0.01$. The estimated α approaches α_{true} as $\hat{r}$ decreases to the true rank $r = 6$. Once the rank is underestimated, the estimated α diverges from the true α. Fig. 3(b) shows the optimal function values for (5.2). The values remain close to 0 while $\hat{r} \geqslant r$, but vary greatly when $\hat{r} < r$. This phenomenon becomes more pronounced as the noise level σ_E decreases. Thus, this could be one clue to choosing an appropriate rank r when the noise level is low.

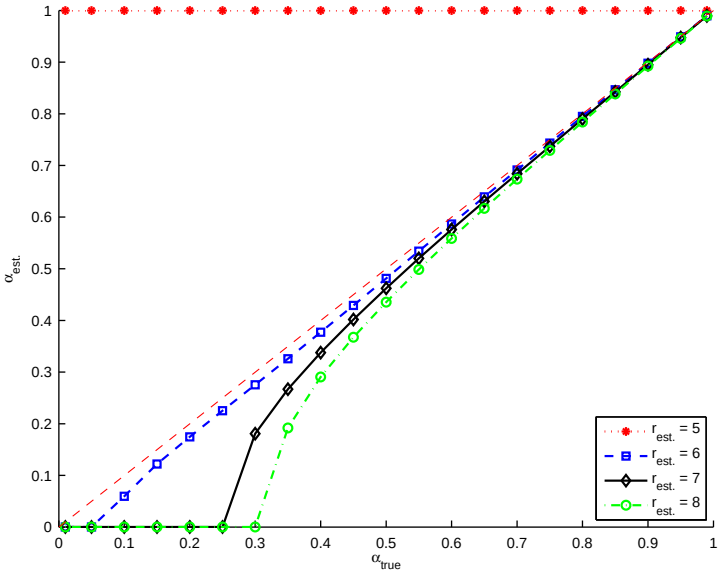
Fig. 4(a) shows the corresponding results for (5.3) when σ_B is known. The α estimation is even more stable than in the previous case when $\hat{r}$ is overestimated. Interestingly, when $\hat{r}$ is underestimated, so is α (red dotted line). Fig. 4(b) represents the ratio of the estimated $\alpha(\hat{r})$ to the estimated $\alpha(\hat{r}+1)$. The ratio stays close to 1 while $\hat{r} \geqslant r$, but is much smaller when $\hat{r} < r$. Even when the noise level is relatively high ($\sigma_E = 0.01$), this decrease is distinguishable, but it is larger as the noise level decreases. Therefore, this ratio of the minimizers α could be an alternative criterion to determine the rank r.

Fig. 5(a) shows the estimated α based on (5.4), used when neither σ_A nor σ_B is known. Similar to the previous cases, the estimated α approaches the true α as $\hat{r}$ approaches the true rank r from above, but the estimation of α fails when $\hat{r} < r$. Fig. 5(b) shows the minimized coefficient of variation, for different noise levels. While the minimized dispersion remains close to zero when $\hat{r} \geqslant r$, the dispersion jumps to a large value (greater than 0.5) when $\hat{r} < r$. The jump becomes more prominent as the noise level decreases. Hence, this is another criterion to determine the rank r. Extensive experiments revealed that rank-determination using c_v is more reliable than the other methods. Since it requires no prior information about σ_A and σ_B, we recommend using this rank-determining strategy to confirm the rank determined by other methods, whenever $n + d - r > 1$.

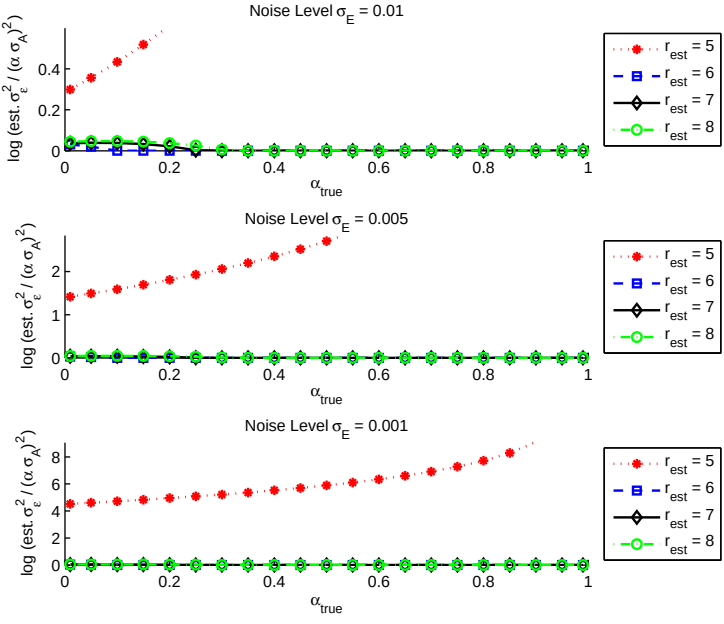
Next we examine the effect of sample size $(n + d - r)$ in the (5.4) method. We may suspect that the dispersion measure may not be reliable if $n + d - r$ is too small, so we set $m = 200, n = 8, r = 7$, and vary d from 1 to 5. Fig. 6 shows the estimated α for different values of d. As d increases, the estimate tends to improve, but it is generally good (for moderately large values of α) even for small $n + d - r$.

Finally, we test how the number of observations m affects the estimation of α. Since all of our methods are based on an asymptotic property of the smallest singular values, we expect that increasing

[3] Even though we generate normally-distributed errors $\boldsymbol{E}_A$ and $\boldsymbol{E}_B$ for the experiments, our methods are not restricted to a particular distribution as long as the errors are uncorrelated with identical variances.

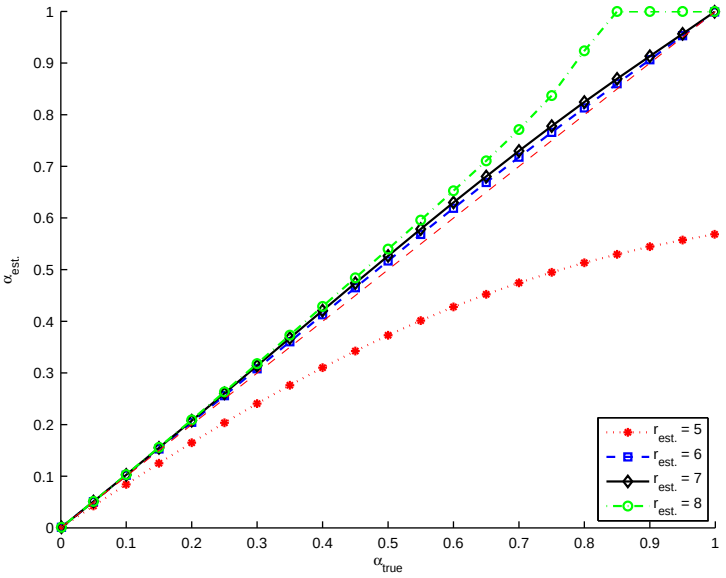


(a) Estimated α vs. α_{true} using (5.2), with noise level $\sigma_E = 0.01$.

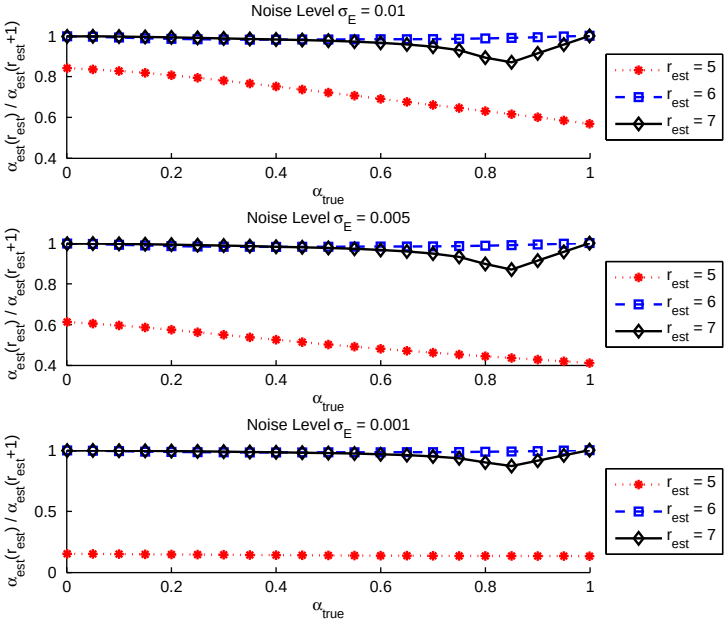


(b) Function value from (5.2)

Fig. 3. Results when σ_A is known: $m = 200, n = 8, d = 4, r = 6$. (a) Estimated α vs. α_{true} using (5.2), with noise level $\sigma_E = 0.01$. (b) Function value from (5.2).

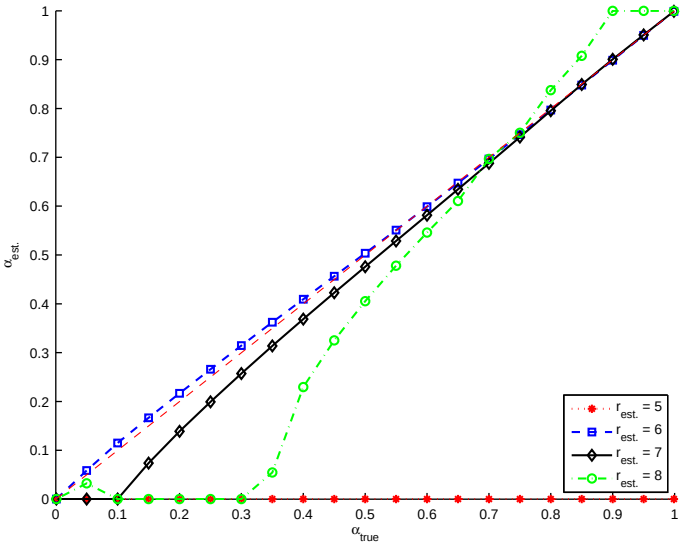


(a) Estimated α vs. α_{true} using (5.3), with noise level $\sigma_E = 0.01$.

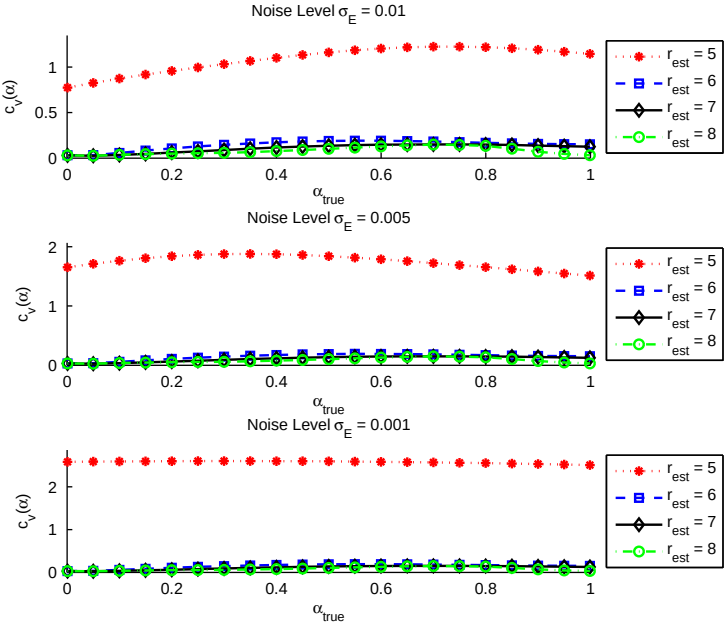


(b) Function value from (5.3)

Fig. 4. Results when σ_B is known: $m = 200, n = 8, d = 4, r = 6$. (a) Estimated α vs. α_{true} using (5.3), with noise level $\sigma_E = 0.01$. (b) Stability of α estimates.



(a) The estimated α vs. α_{true} for different rank estimates $\widehat{r}$, with noise level $\sigma_E = 0.01$.



(b) The estimated coefficient of variation, c_v, for different choices of $\widehat{r}$ and noise level.

Fig. 5. The result from using (5.4) to determine α for $m = 200$, $n = 8$, $d = 4$, $r = 6$. (a) The estimated α vs. α_{true} for different rank estimates $\hat{r}$, with noise level $\sigma_E = 0.01$. (b) The estimated coefficient of variation, c_v, for different choices of $\hat{r}$ and noise level.

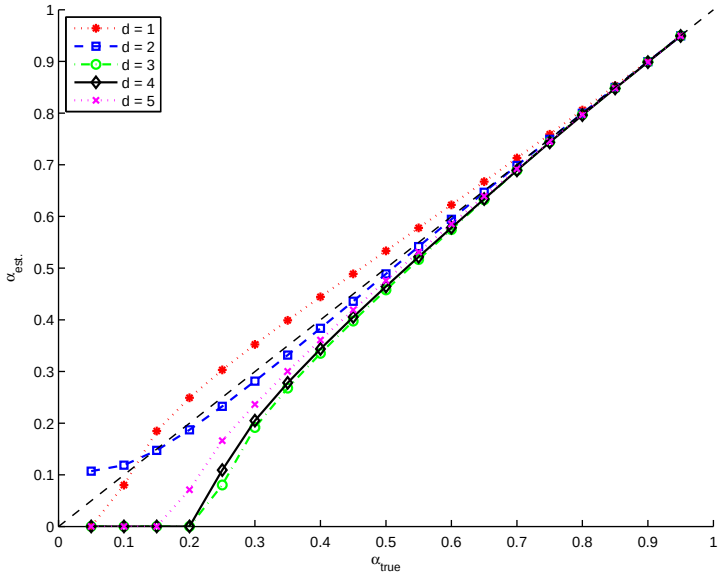


Fig. 6. Estimated α vs. α_{true} using the coefficient of variation for $m = 200$, $n = 8$, and $r = 7$, $\sigma_E = 0.01$, varying the number of right-hand sides d from 1 to 5.

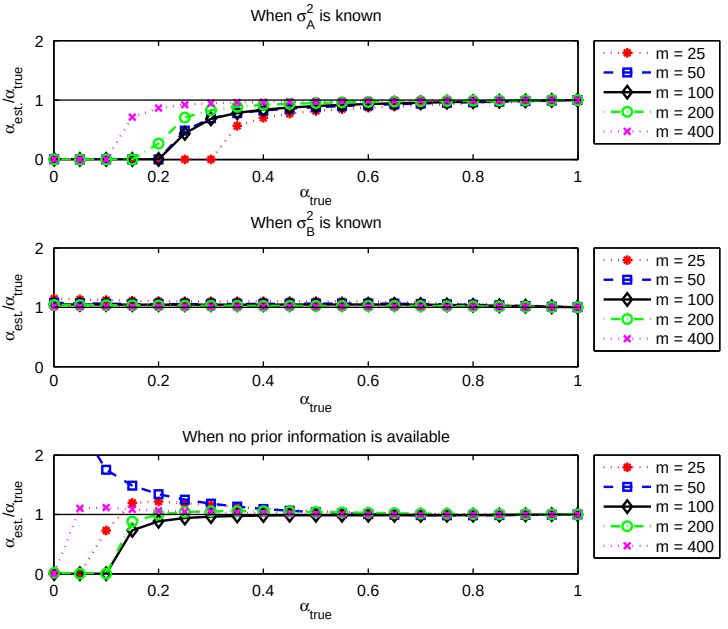


Fig. 7. Known rank: α-ratio for $n = 8$, $r = 6$, and $d = 4$, $\sigma_E = 0.01$, varying the number of observations m.

m should improve the quality of the estimate of α. Fig. 7 shows the relative error in the α estimates as m varies between 25 and 400. The estimation does improve with larger m for all proposed methods, and estimation by (5.3) (with a known σ_B^2) shows the most reliable performance even with small m.

7. Discussion and conclusions

We have defined an implicitly-weighted TLS formulation (ITLS) that includes LS, TLS, and DLS as special cases as a parameter varies between 0 and 1. We have discussed the role of the ratio between the variances of errors in $\boldsymbol{A}$ and $\boldsymbol{B}$ in choosing an appropriate parameter in ITLS. We derived asymptotic properties of the estimate as the number of observations $m \to \infty$, even when the model is rank deficient. We also proposed methods for computing the family of solutions efficiently. We developed algorithms for choosing the appropriate solution when only σ_A^2 or σ_B^2 is known, or neither is known, in which case we require $n + d - r > 1$. We provided experimental results on the usefulness of the ITLS (or, equivalently the STLS) family of solutions, and on our algorithms for estimating α and r.

It would be easy to add a regularization term to the ITLS problem, in order to handle discrete ill-posed problems.

This work leaves two important open questions. First, the concept of a *core problem* [7], so useful for a single right-hand side, does not completely explain the character of TLS problems when $d > 1$ [8], and more work is needed. This is related to the choice of $\hat{t}$. Second, our parameter choice algorithm requires an estimate of either σ_A or σ_B when $n + d - r = 1$, a single right-hand-side problem with full rank, so more work on that case is needed.

Acknowledgement

We are grateful to the referees for helpful comments and queries.

References

[1] Richard P. Brent, Algorithms for Minimization without Derivatives, Prentice-Hall, 1973.
[2] T.F. Chan, An improved algorithm for computing the singular value decomposition, ACM Trans. Math. Software 80 (1982) 72–78.
[3] R.D. Fierro, L. Vanhamme, S. Van Huffel, Total least squares algorithms based on rank-revealing complete orthogonal decompositions, in: Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1997, pp. 99–116.
[4] Leon Jay Gleser, Estimation in a multivariate "errors in variables" regression model: large sample results, Ann. Statist. 9 (1) (1981) 24–44.
[5] G.H. Golub, W. Kahan, Calculating the singular values and pseudo-inverse of a matrix, SIAM J. Ser. B, Numer. Anal. 2 (1965) 205–224.
[6] R.B. Lehoucq, D.C. Sorensen, C. Yang, ARPACK Users Guide: Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods, SIAM, Philadelphia, 1998.
[7] Christopher C. Paige, Zdeněk Strakoš, Scaled total least squares fundamentals, Numer. Math. 91 (2002) 117–146.
[8] Martin Plešinger, The Total Least Squares Problem and Reduction of Data in $AX \approx B$. Ph.D. Thesis, Technical University of Liberec, Liberec, Czech Republic, March 2008.
[9] Bhaskar D. Rao, Unified treatment of LS, TLS and truncated SVD methods using a weighted TLS framework, in: Sabine Van Huffel (Ed.), Recent Advances in Total Least Squares Techniques and Errors-in-Variables Modeling, SIAM, 1997, pp. 11–20.
[10] G.W. Stewart, Updating a rank-revealing ULV decomposition, SIAM J. Matrix Anal. Appl. 14 (1993) 494–499.
[11] Sabine Van Huffel, Joos Vandewalle, The Total Least Squares Problem: Computational Aspects and Analysis, SIAM, Philadelphia, 1991.